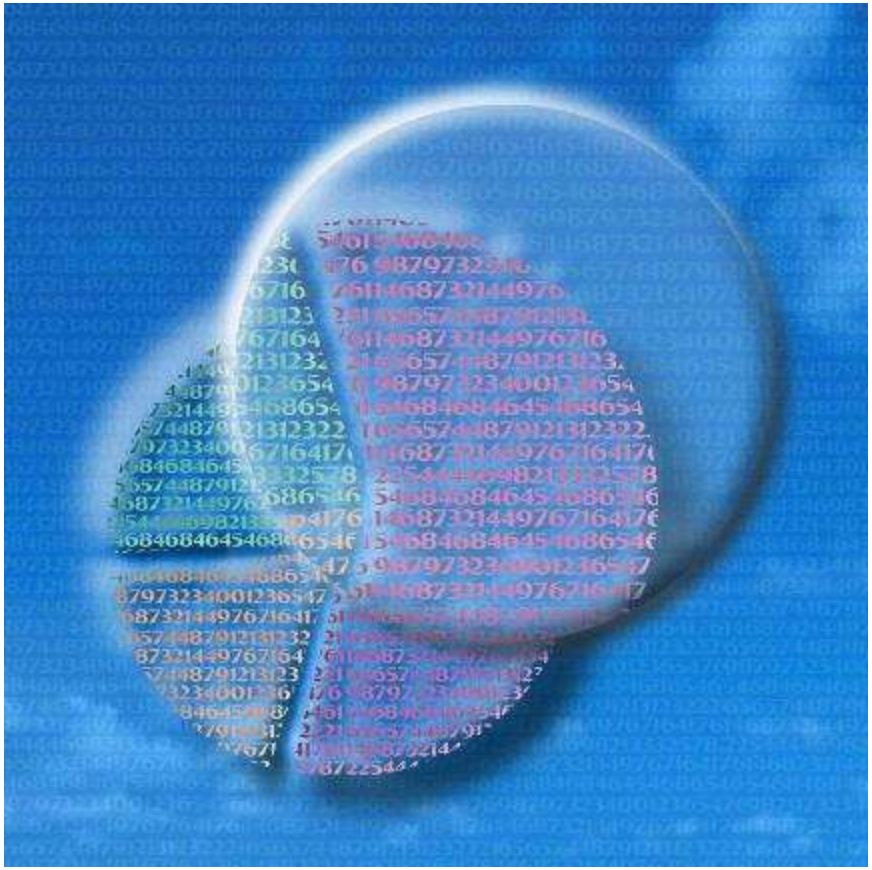


Sensitivity analysis: An introduction

Trondheim, October 10, 2016

Andrea Saltelli

European Centre for Governance in Complexity, Universities of Bergen (NO) and Autonoma of Barcelona (ES), andrea.saltelli@uib.no

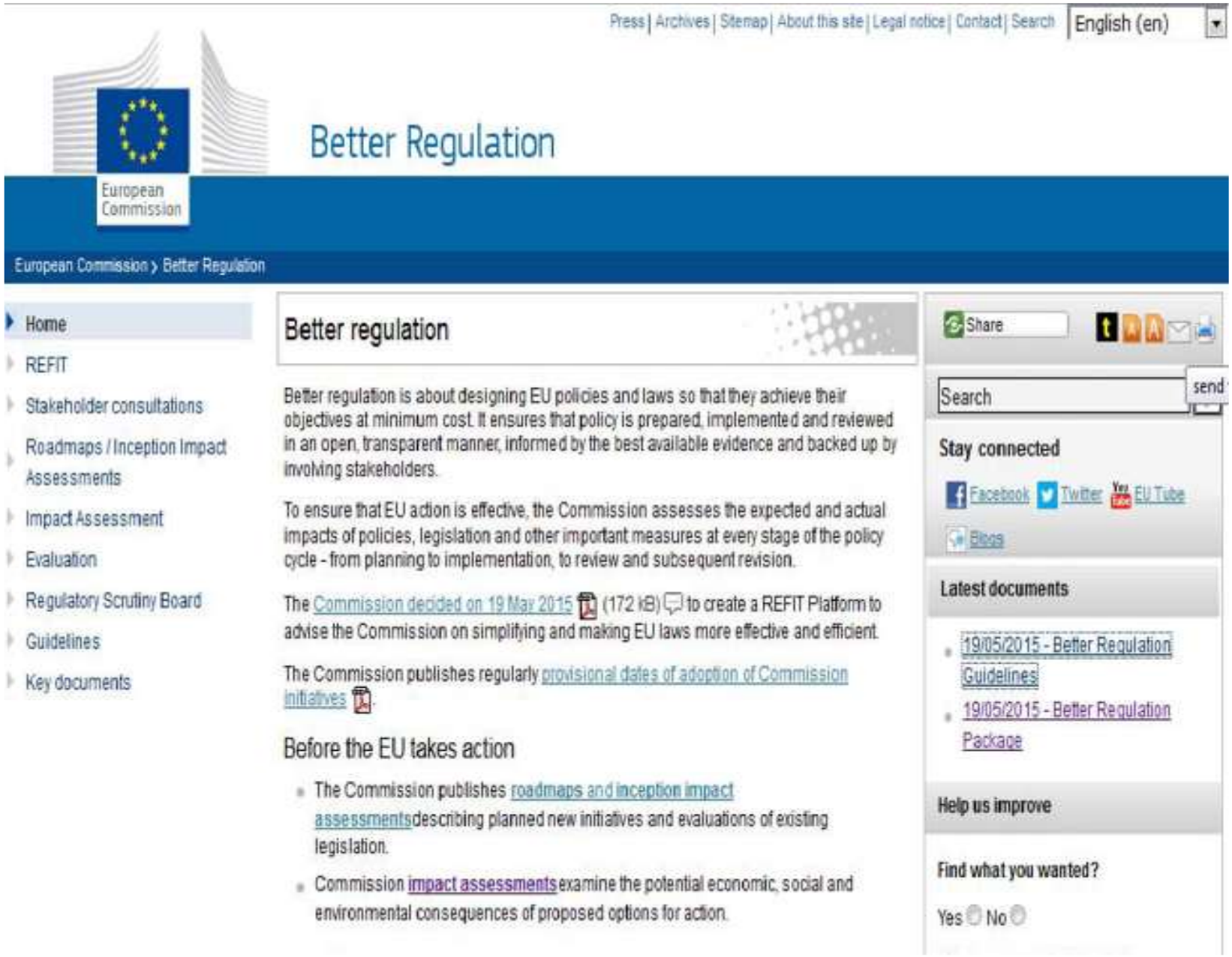
@andreasaltelli

Where to find this talk: www.andreasaltelli.eu

sensitivity analysis, sensitivity auditing, science for policy, impact assessment

Why sensitivity analysis

http://ec.europa.eu/smart-regulation/



http://ec.europa.eu/smart-regulation/guidelines/docs/br_toolbox_en.pdf

When testing the evidence behind inference some reasonable people suggest that ‘sensitivity analysis would help’

...

Edward E. Leamer, 1990, Let's Take the Con Out of Econometrics, American Economics Review, 73 (March 1983), 31-43.

<<I have proposed a form of organised sensitivity analysis that I call “global sensitivity analysis” in which a neighborhood of alternative assumptions is selected and the corresponding interval of inferences is identified.>>

Edward E. Leamer, 1990, Let's Take the Con Out of Econometrics, American Economics Review, 73 (March 1983), 31-43.

<<Conclusions are judged to be sturdy only if the neighborhood of assumptions is wide enough to be credible and the corresponding interval of inferences is narrow enough to be useful.>>

Funtowicz & Ravetz's GIGO (Garbage In, Garbage Out) Science – or pseudo-science – "where uncertainties in inputs must be suppressed least outputs become indeterminate"

Leamer's 'Conclusions are judged to be sturdy only if the neighborhood of assumptions is wide enough to be credible and the corresponding interval of inferences is narrow enough to be useful'.

The definition of pseudo-science from the 1990 book of Silvio Funtowicz & Jerome R. Ravetz's implies some form uncertainty analysis.

Funtowicz, S. O. and Ravetz, J. R., 1990. Uncertainty and quality in science for policy. Dordrecht: Kluwer.

Back to Leamer:
"With the ashes of the mathematical models used to rate mortgage-backed securities still smoldering on Wall Street, now is an ideal time to revisit the sensitivity issues"

Tantalus on the Road to Asymptopia
Edward E. Leamer, 2010 *Journal of Economic Perspectives*, **24**, (2), 31–46.

"… my observation of economists at work who routinely pass their data through the filters of many models and then choose a few results for reporting purposes."

Ibidem

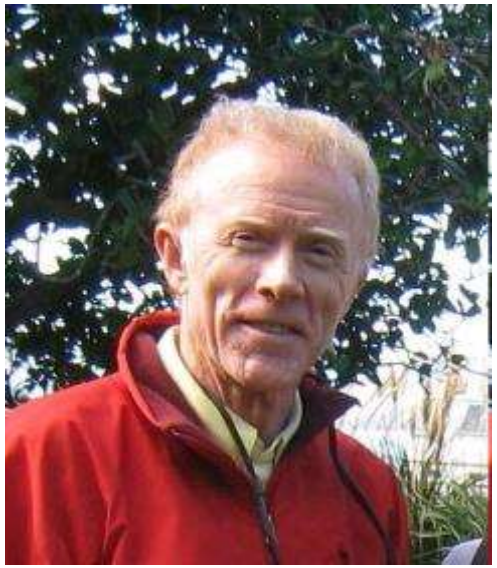
"One reason these methods are rarely used is their honesty seems destructive;"

Ibidem

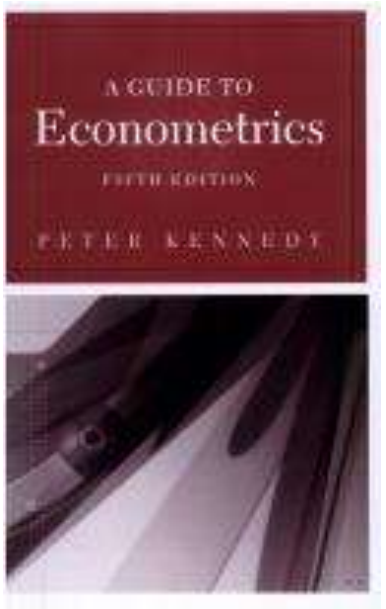
"or, to put it another way, a fanatical commitment to fanciful formal models is often needed to create the appearance of progress."

Ibidem

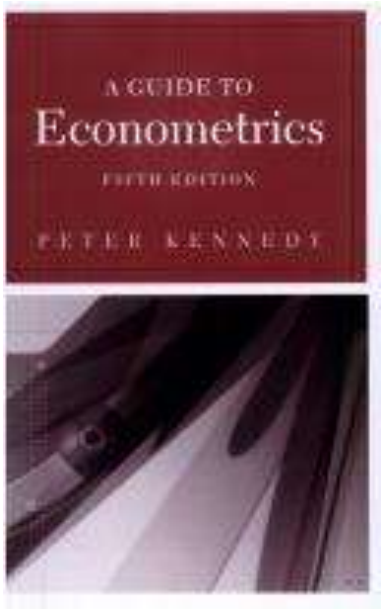
Peter Kennedy, A Guide to Econometrics.
Anticipating criticism by applying sensitivity analysis. This is one of the ten commandments of applied econometrics according to Peter Kennedy:



<<Thou shall confess in the presence of sensitivity.
Corollary: Thou shall anticipate criticism >>



<<When reporting a sensitivity analysis, researchers should explain fully their specification search so that the readers can judge for themselves how the results may have been affected. This is basically an `honesty is the best policy' approach, […]'.>>



Today: from p-hacking to the 'Mathiness' discussion: blogs of Paul Romer, Judith Curry; Erik Reinert's 'scholasticism' paper.

See https://paulromer.net/mathiness/

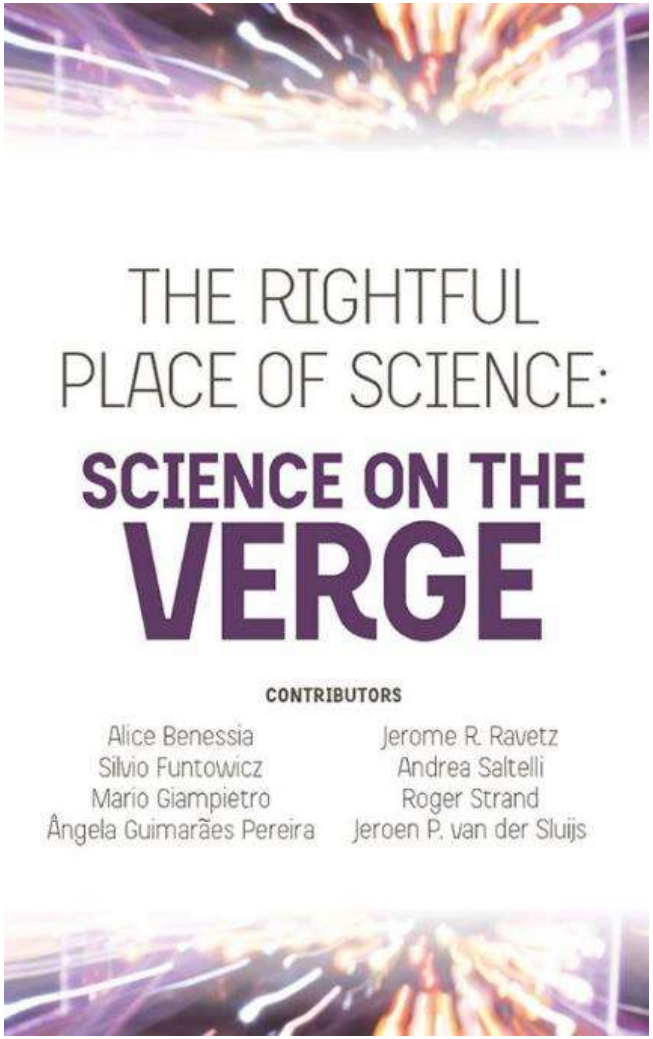
https://judithcurry.com/2015/08/12/the-adversarial-method-versus-feynman-integrity-2/

http://www.andreasaltelli.eu/file/repository/Full_Circle_scholasticism_2.pdf

Paul Romer

Judith Curry

Erik Reinert



The Rightful Place of Science: Science on the Verge

Paperback – 20 Feb 2016

by Andrea Saltelli (Author), Alice Benessia (Author), & 7 more

1 customer review

See all formats and editions

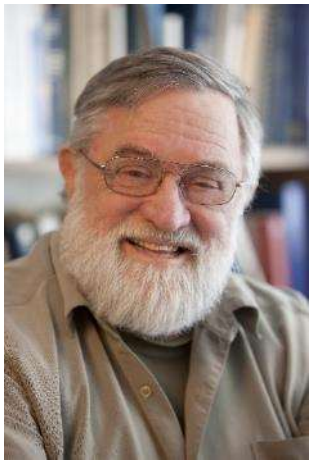
Kindle Edition	Paperback
£3.61	£6.99

More on these later today

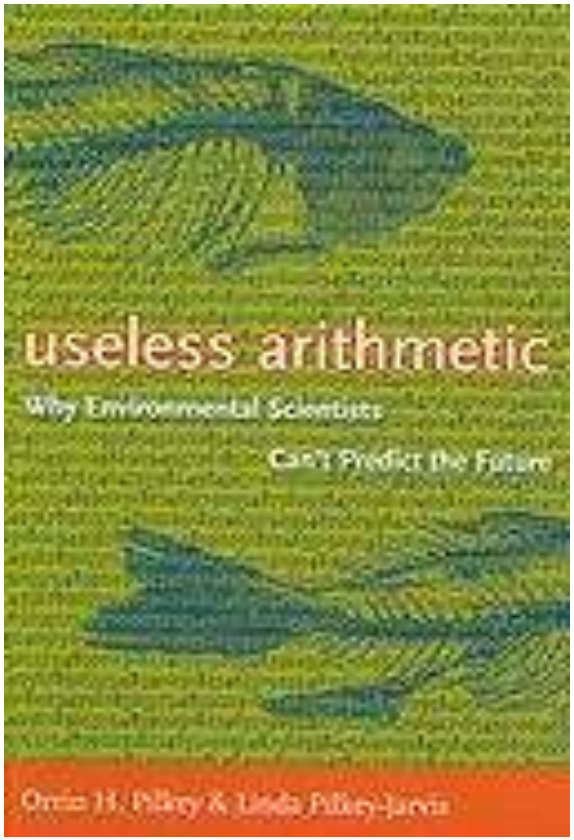
http://www.amazon.com/Rightful-Place-Science-Verge/dp/0692596380/ref=sr_1_1?s=books&ie=UTF8&qid=1456255907&sr=1-1&keywords=saltelli

http://www.andreasaltelli.eu/science-on-the-verge

Limits of sensitivity analysis



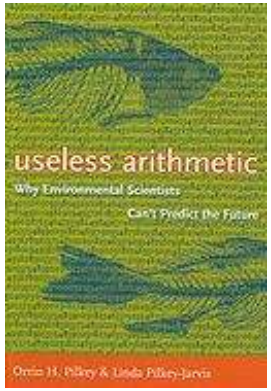
Orrin H. Pilkey
Duke University, NC



Useless Arithmetic: Why Environmental Scientists Can't Predict the Future

by Orrin H. Pilkey and Linda Pilkey-Jarvis

'Quantitative mathematical models used by policy makers and government administrators to form environmental policies are seriously flawed'

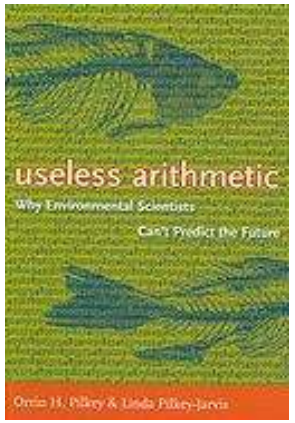


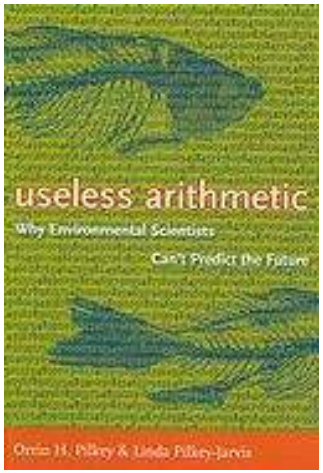
<<It is important, however, to recognize that the sensitivity of the parameter in the equation is what is being determined, not the sensitivity of the parameter in nature.

[…] If the model is wrong or if it is a poor representation of reality, determining the sensitivity of an individual parameter in the model is a meaningless pursuit.>>

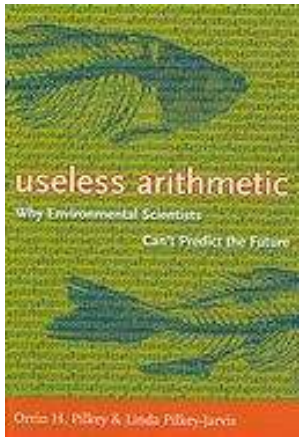
One of the examples discussed concerns the Yucca Mountain repository for radioactive waste. TSPA model (for total system performance assessment) for safety analysis.

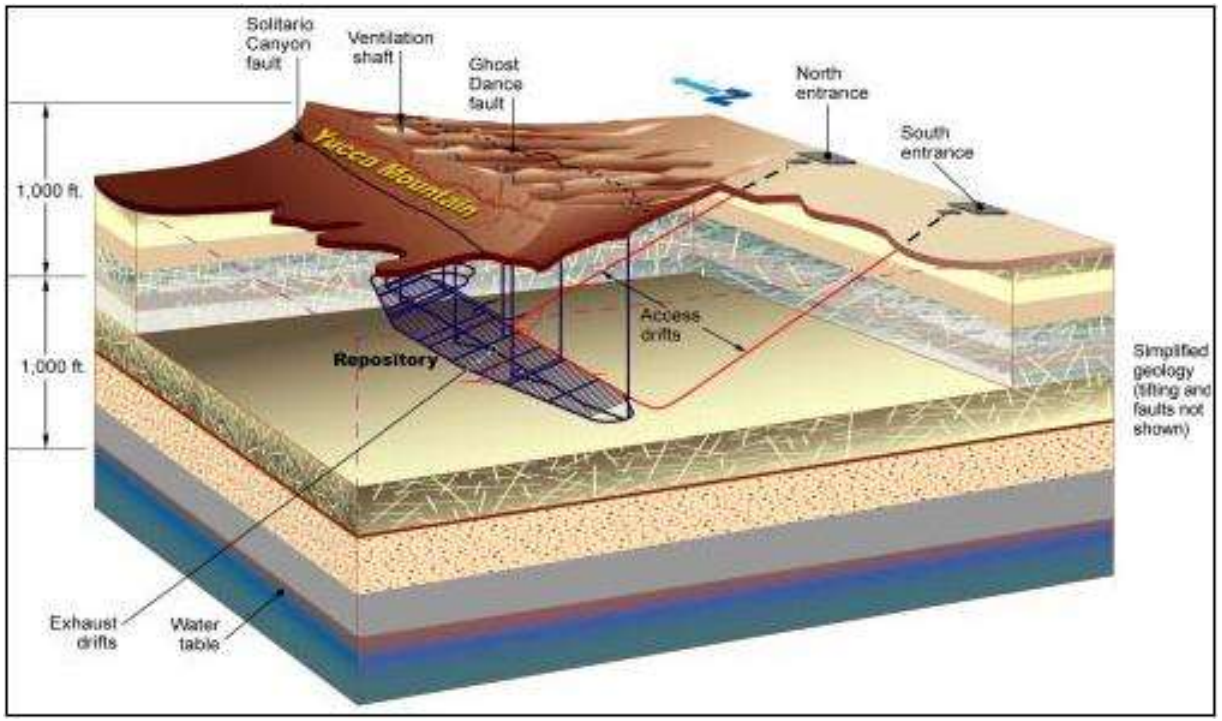
TSPA is Composed of 286 sub-models.





TSPA (like any other model) relies on assumptions → one is the low permeability of the geological formation → long time for the water to percolate from surface to disposal.



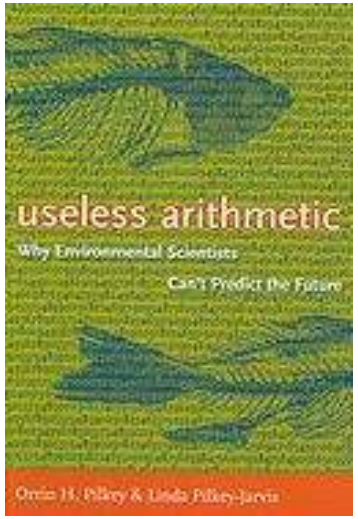


The confidence of the stakeholders in TSPA was not helped when evidence was produced which could lead to an upward revision of 4 orders of magnitude of this parameter (the ^{36}Cl story)

Type III error in sensitivity: Examples:

In the case of TSPA (Yucca mountain) a range of 0.02 to 1 millimetre per year was used for percolation of flux rate.

→… SA useless if it is instead ~ 3,000 millimetres per year.



"Scientific mathematical modelling should involve constant efforts to falsify the model"

Ref. ➔ Robert K. Merton's 'Organized skepticism '

Communalism – the common ownership of scient40

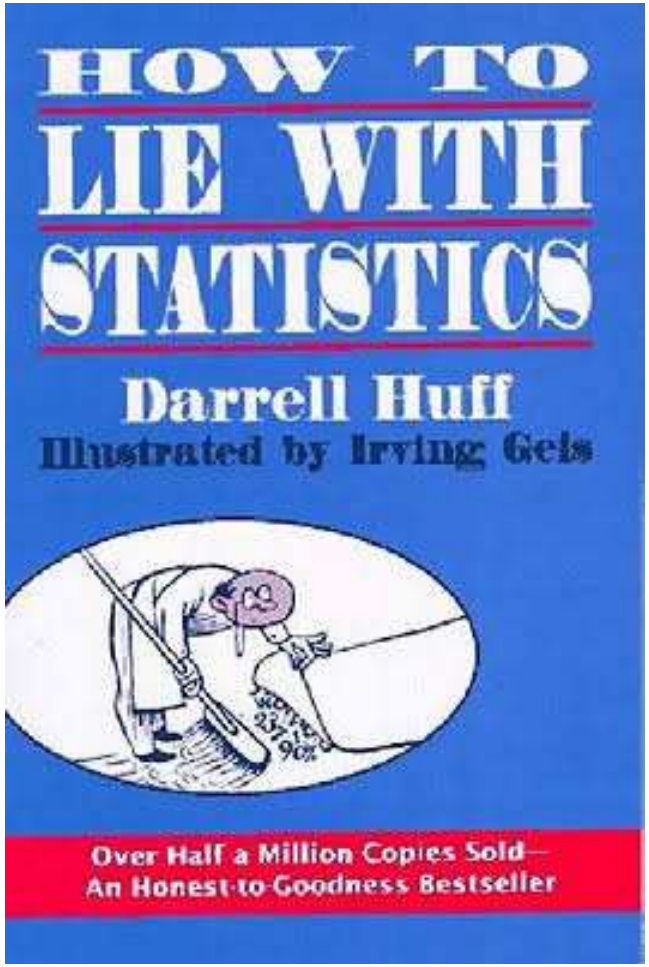
ific discoveries, according to which scientists give up intellectual property rights in exchange for recognition and esteem (Merton actually used the term Communism, but had this notion of communalism in mind, not Marxism);

Universalism – according to which claims to truth are evaluated in terms of universal or impersonal criteria, and not on the basis of race, class, gender, religion, or nationality;

Disinterestedness – according to which scientists are rewarded for acting in ways that outwardly appear to be selfless;

Organized Skepticism – all ideas must be tested and are subject to rigorous, structured community scrutiny.

Will any sensitivity analysis do the job? Can I lie with sensitivity analysis as I can lie with statistics?



Saltelli, A., Annoni P., 2010, How to avoid a perfunctory sensitivity analysis, *Environmental Modeling and Software,* **25**, 1508–1517.

What do these have in common?

J. Campbell, *et al.*, *Science* **322**, 1085 (2008).
R. Bailis, M. Ezzati, D. Kammen, *Science* **308**, 98 (2005).
E. Stites, P. Trampont, Z. Ma, K. Ravichandran, *Science* **318**, 463 (2007).
J. Murphy, et al., *Nature* **430**, 768-772 (2004).
J. Coggan, *et al.*, *Science* **309**, 446 (2005).

They use a one factor at a time approach (OAT)

OAT methods – derivatives – local

$$S_1 = \frac{\partial Y}{\partial x_j}$$

Effect on Y of perturbing x_j around its nominal value x_j^0

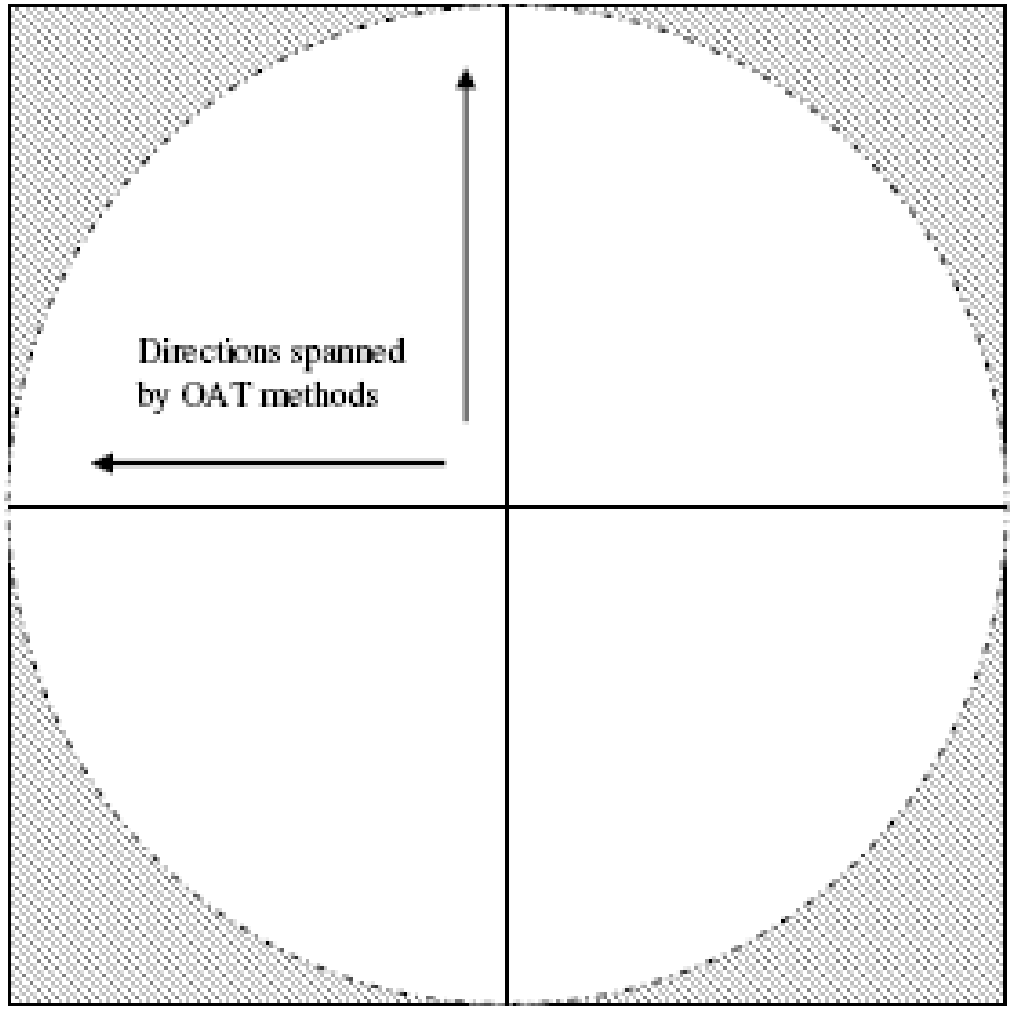
$$S_2 = \frac{\partial Y}{\partial x_j}\frac{x_j^0}{Y^0}$$

Relative effect on Y of perturbing x_j by a fixed fraction of its nominal value x_j^0

$$S_3 = \frac{\partial Y}{\partial x_j}\frac{std(x_j)}{std(Y)}$$

Relative effect on Y of perturbing x_j by a fixed fraction of its standard deviation

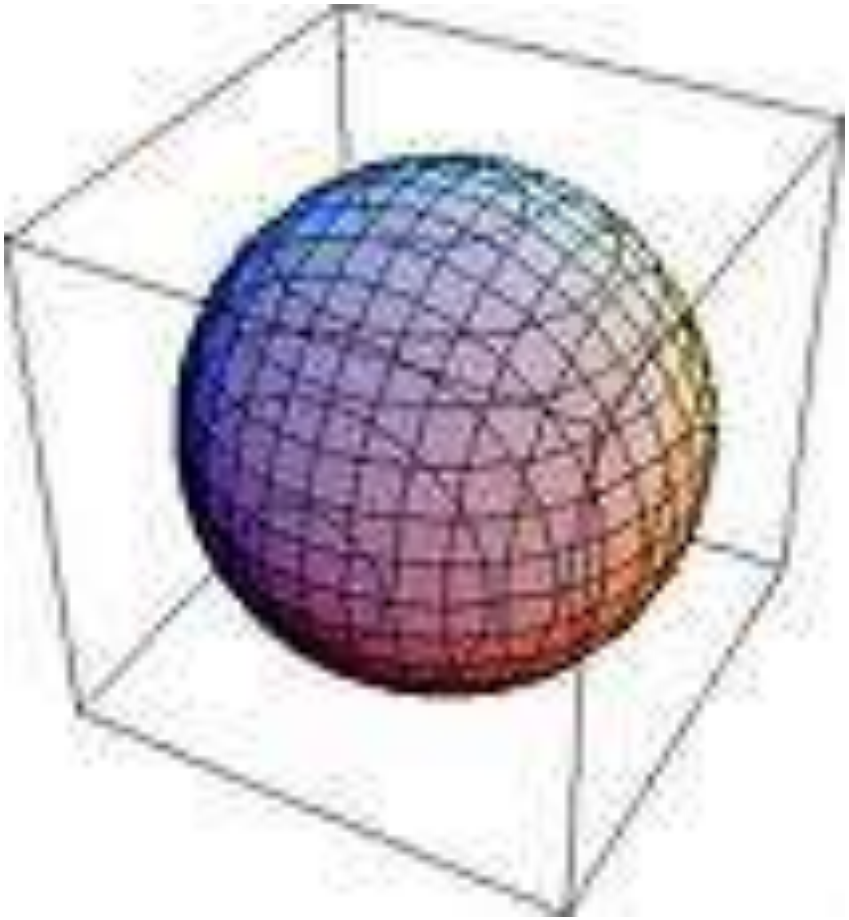
OAT in 2 dimensions



Area circle / area square =?

~ 3/4

OAT in 3 dimensions



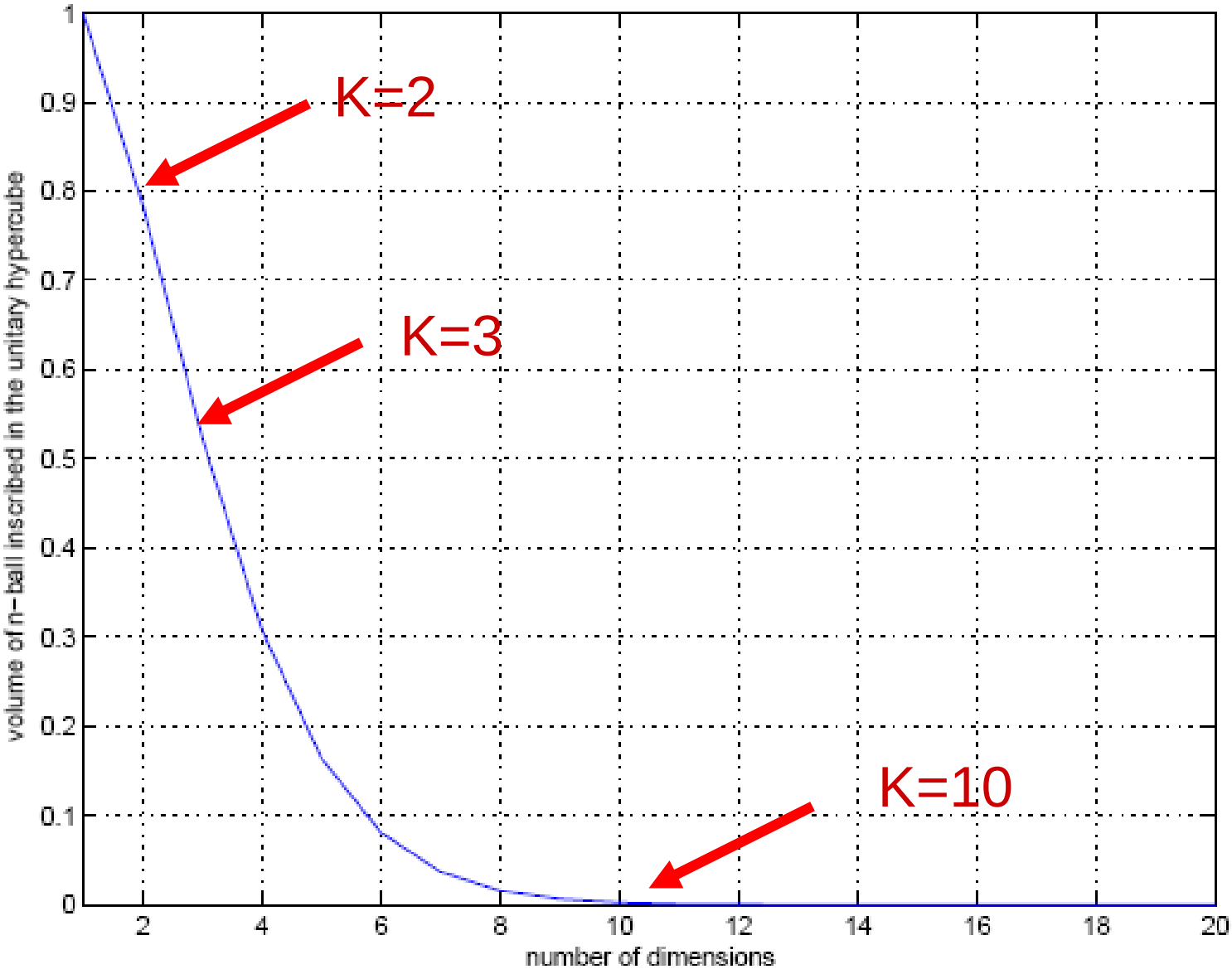
Volume sphere / volume cube =?

~ 1/2

OAT in 10 dimensions

Volume hypersphere / volume ten dimensional hypercube ~ 0.0025

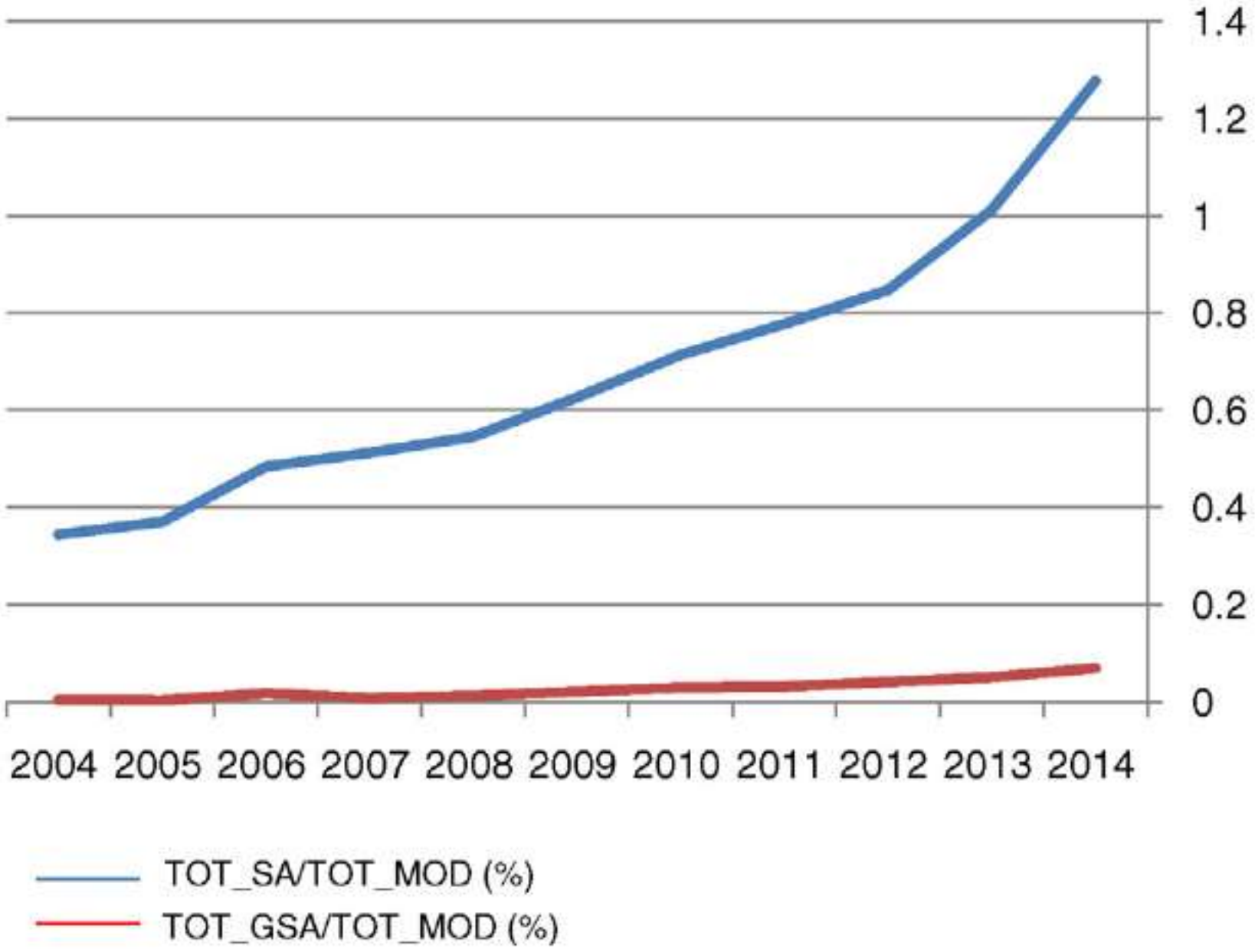
OAT in k dimensions



How are we doing in 2016?

…OAT is still the most largely used technique in SA, … clear increase in the use of GSA with preference for regression and variance-based techniques.

Ferretti, F., Saltelli A., Tarantola, S., 2016, Trends in Sensitivity Analysis practice in the last decade, Science of the Total Environment, http://dx.doi.org/10.1016/j.scitotenv.2016.02.133


1.4
1.2
1
0.8
0.6
0.4
0.2
0
2004 2005 2006 2007 2008 2009 2010 2011 2012 2013 2014
TOT_SA/TOT_MOD (%)
TOT_GSA/TOT_MOD (%)

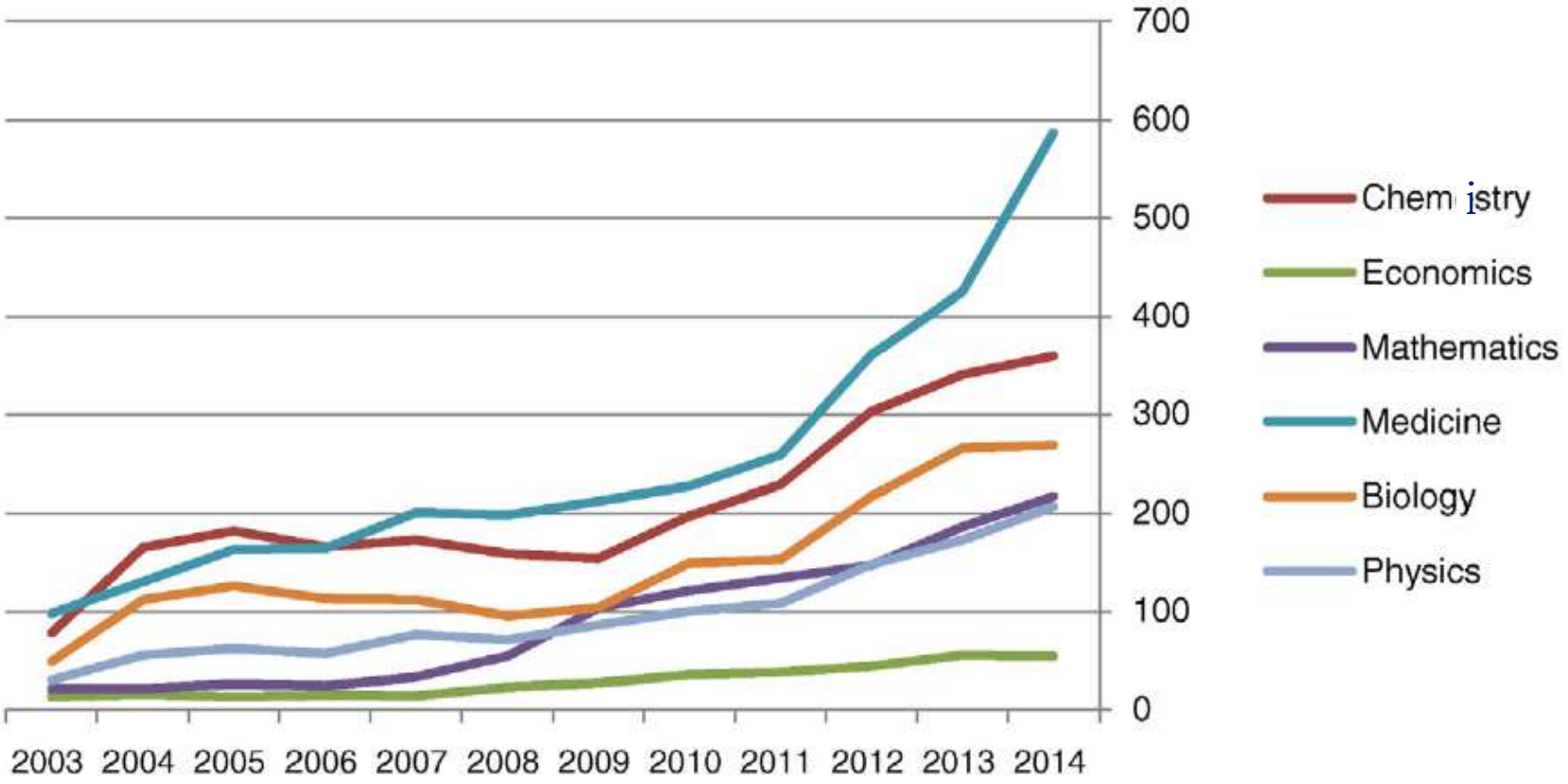


Fig. 4. GSA in the different scientific domains.

Definition of uncertainty and sensitivity analysis.

Sensitivity analysis: The study of the relative importance of different input factors on the model output.

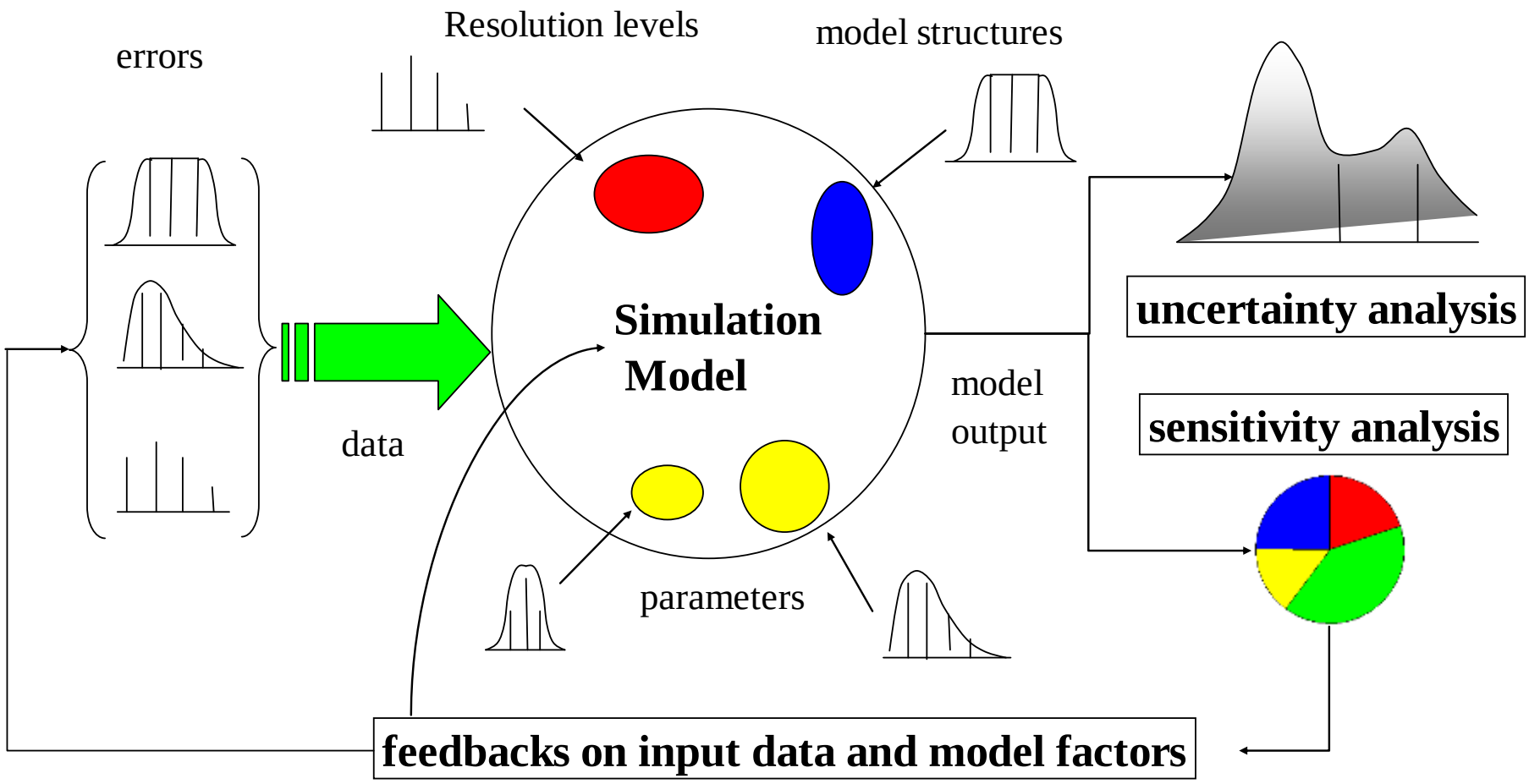
Uncertainty analysis: Focuses on just quantifying the uncertainty in model output.

[Global*] sensitivity analysis: “The study of how the uncertainty in the output of a model (numerical or otherwise) can be apportioned to different sources of uncertainty in the model input”

Saltelli A., 2002, Sensitivity Analysis for Importance Assessment, Risk Analysis, 22 (3), 1-12.

- Modelling in a Monte Carlo framework using quasi MC-points
- All uncertainties activated simultaneously; uncertainty and sensitivity together

An engineer's vision of UA, SA



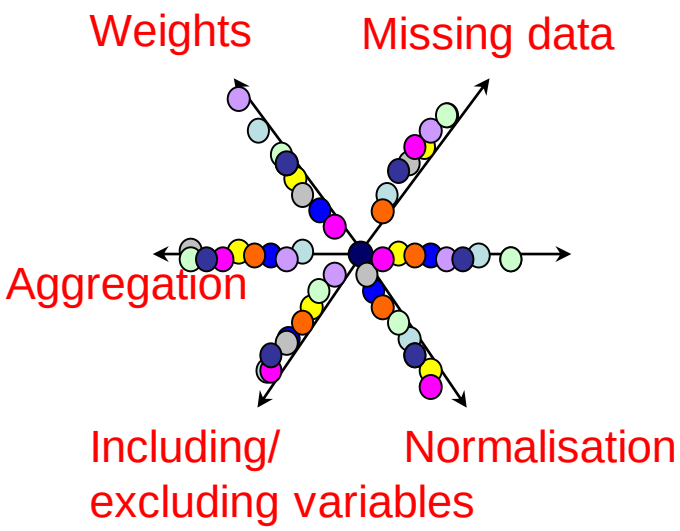
One can sample more than just factors …

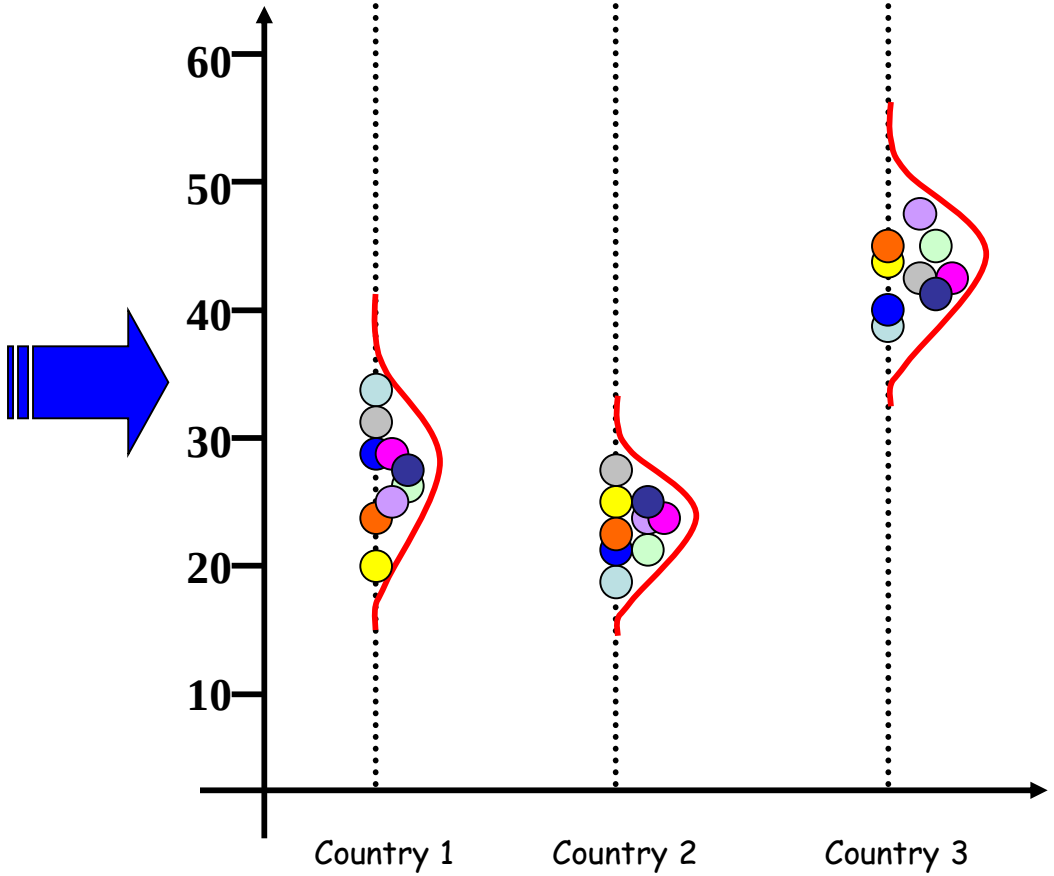
Using triggers one can sample modelling assumptions …

Example: Y is a composite indicator

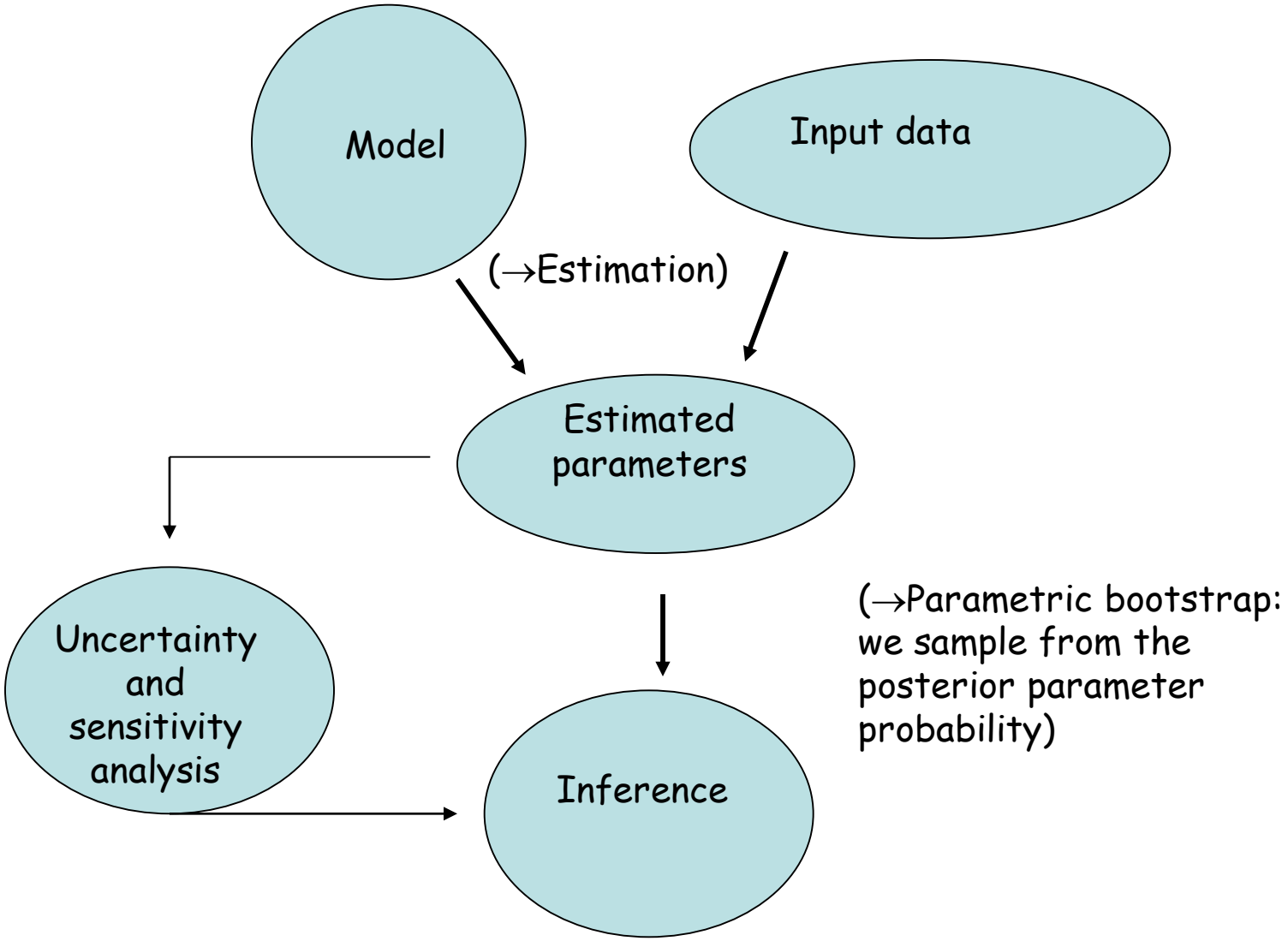
Assumption	**Alternatives**
Number of indicators	▪ all six indicators included or one-at-time excluded (6 options)
Weighting method	▪ original set of weights, ▪ factor analysis, ▪ equal weighting, ▪ data envelopment analysis
Aggregation rule	▪ additive, ▪ multiplicative, ▪ Borda multi-criterion

Space of alternatives





Models maps assumptions onto inferences … (Parametric bootstrap version of UA/SA)



Sample matrix for parametric bootstrap.

$$\begin{matrix} x_{11} & x_{12} & \dots & x_{1k} \\ x_{21} & x_{22} & \dots & x_{2k} \\ \dots & \dots & \dots & \dots \\ x_{N1} & x_{N2} & \dots & x_{Nk} \end{matrix}$$

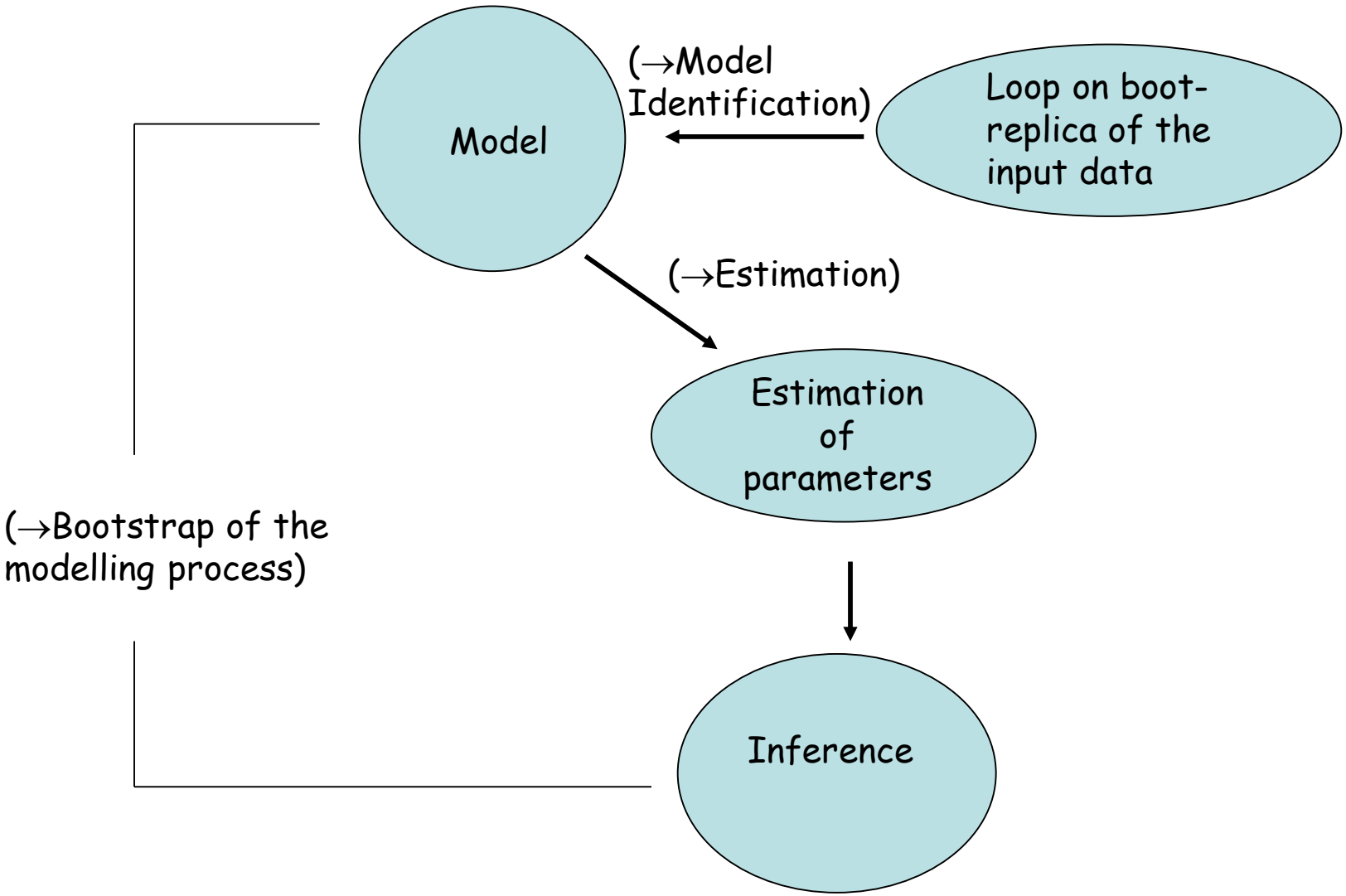
Each row is a sample trial for one model run. Each column is a sample of size N from the marginal distribution of the parameters as generated by the estimation procedure.

Model results:

Each row is the error-free result of the model run.

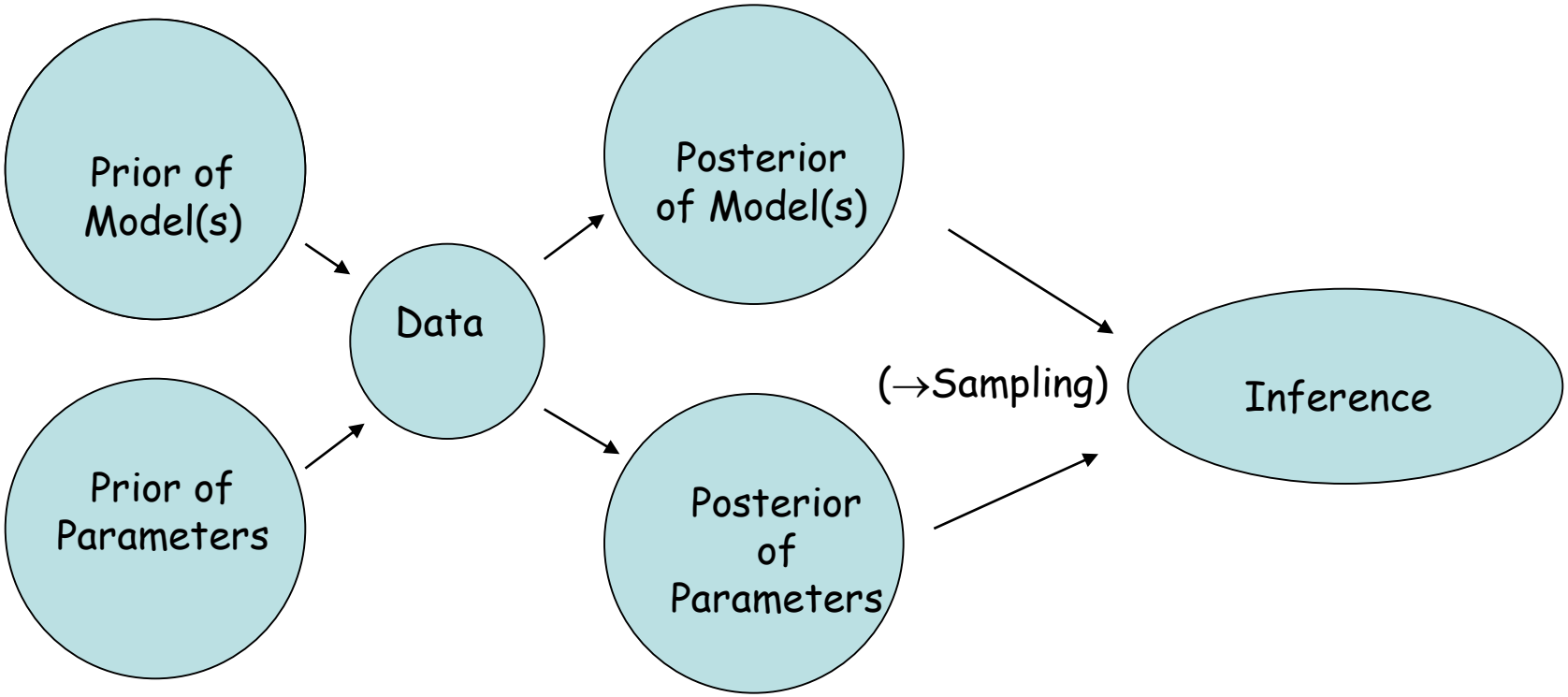
$$y_1$$
$$y_2$$
$$\dots$$
$$y_N$$

Bootstrapping-of-the-modelling-process

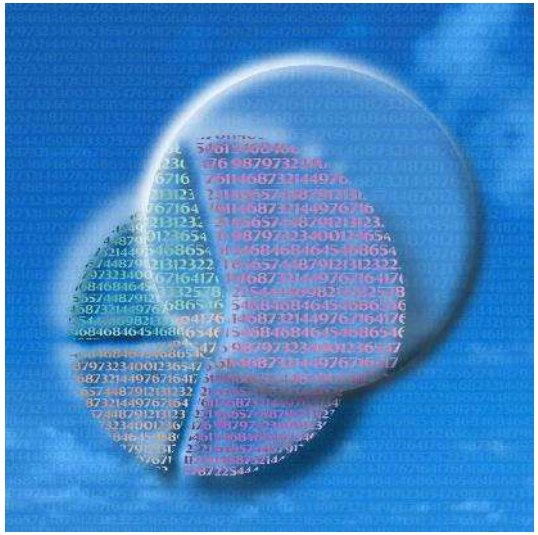


Chatfield, C., 1995, Model Uncertainty, Data Mining and Statistical Inference, Journal of the Royal Statistical Society. Series A (Statistics in Society), 158, No. 3, 419-466.

Bayesian Model Averaging



Hoeting, J.A., Madigan, D., Raftery, A.E. and Volinsky, C.T., 1999, Bayesian Model Averaging: A Tutorial Statistical Science, 1999, Vol. 14, No. 4, 382–417

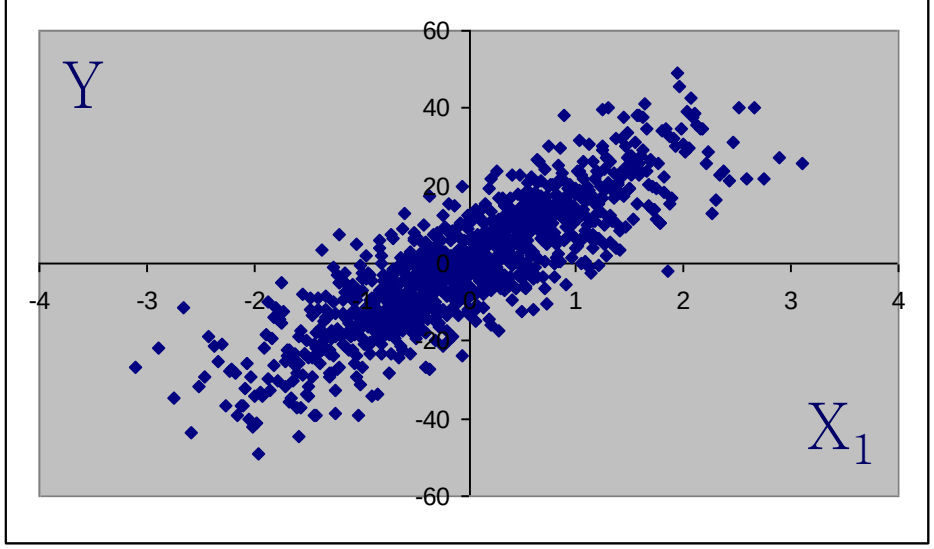


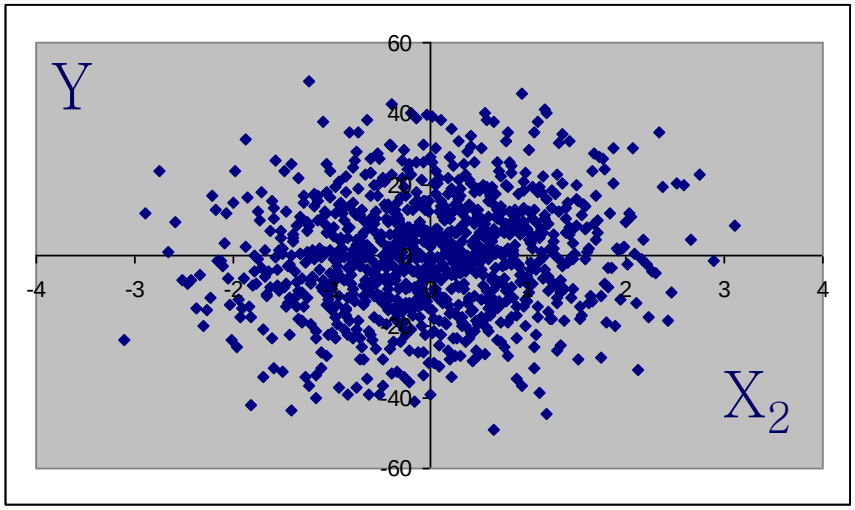
Our preferred methods for SA: variance based

An intuitive derivation of sensitivity indices

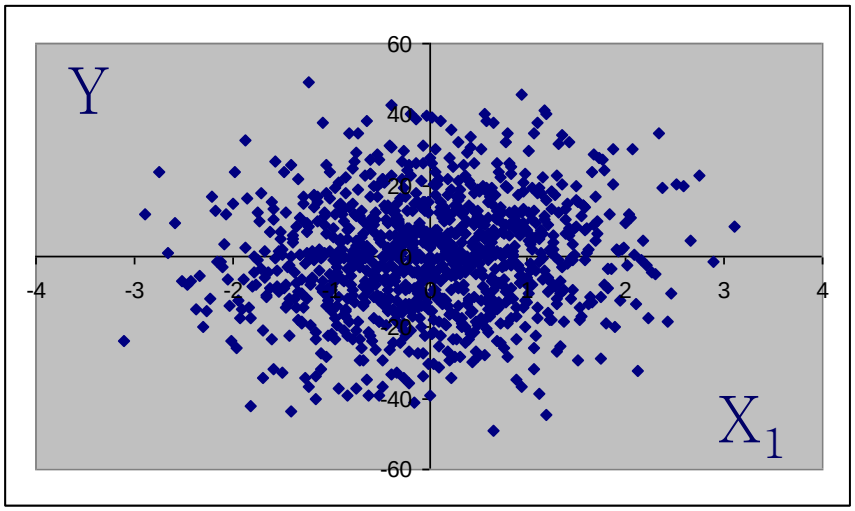
$$\begin{matrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2k} \\ \cdots & \cdots & \cdots & \cdots \\ x_{N1} & x_{N2} & \cdots & x_{Nk} \end{matrix}$$

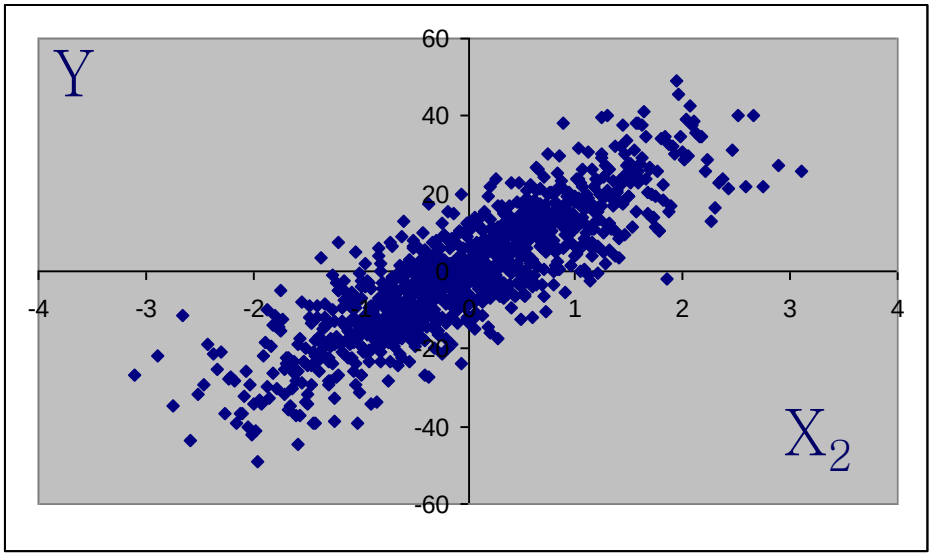
$$\begin{matrix} y_1 \\ y_2 \\ \cdots \\ y_N \end{matrix}$$





Scatterplots of y versus sorted factors

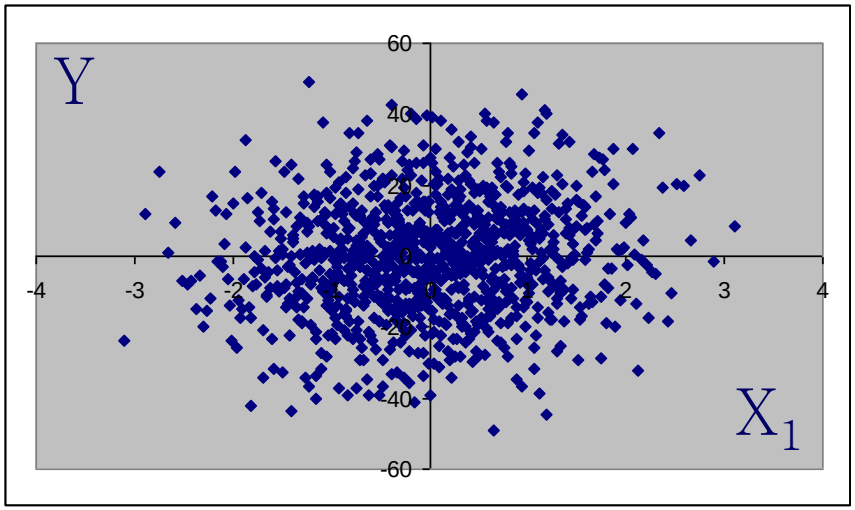


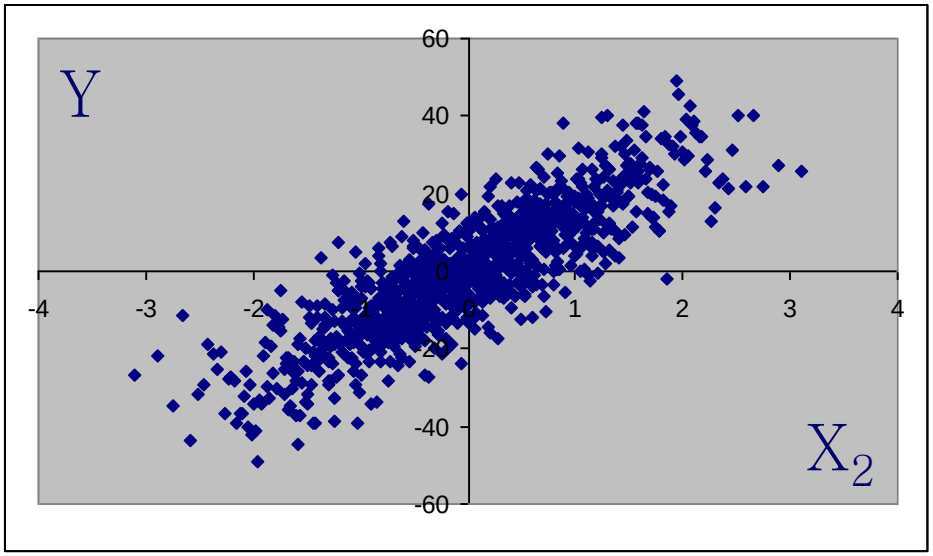


The ordinate axis is always Y

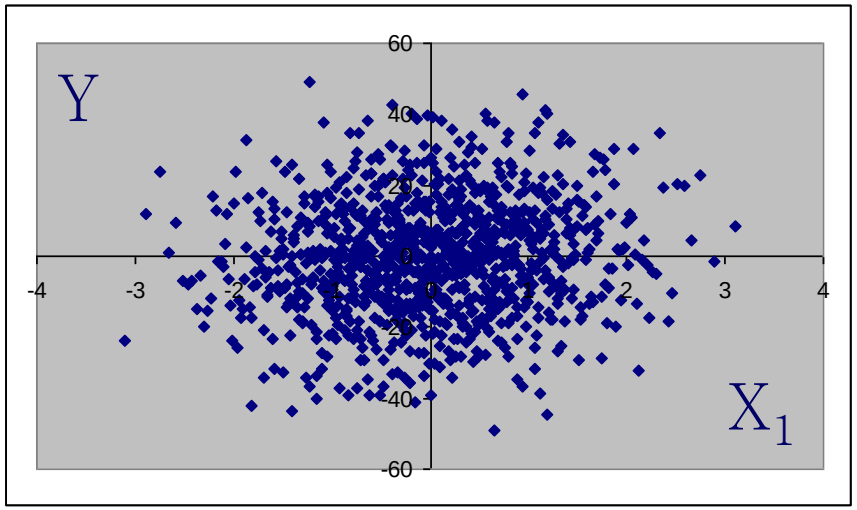
The abscissa are the various factors X_i in turn.

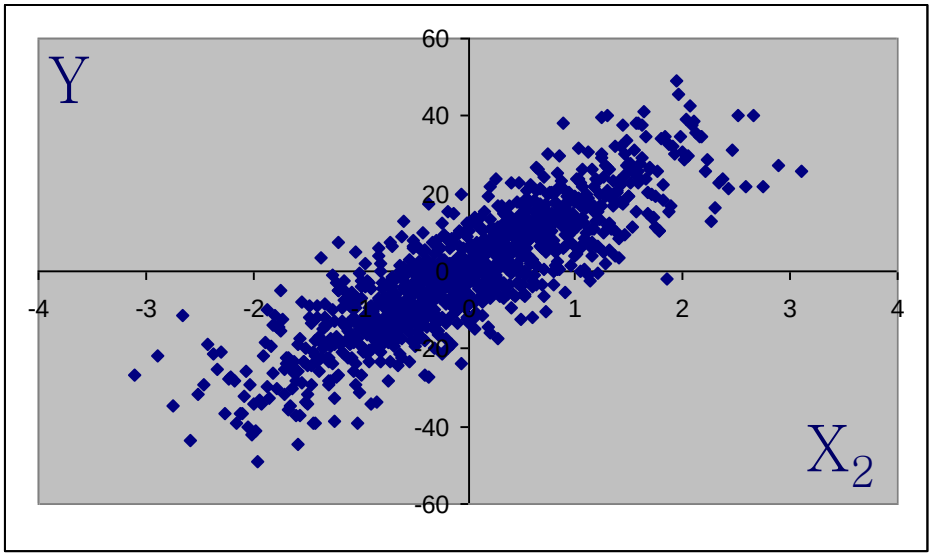
The points are always the same





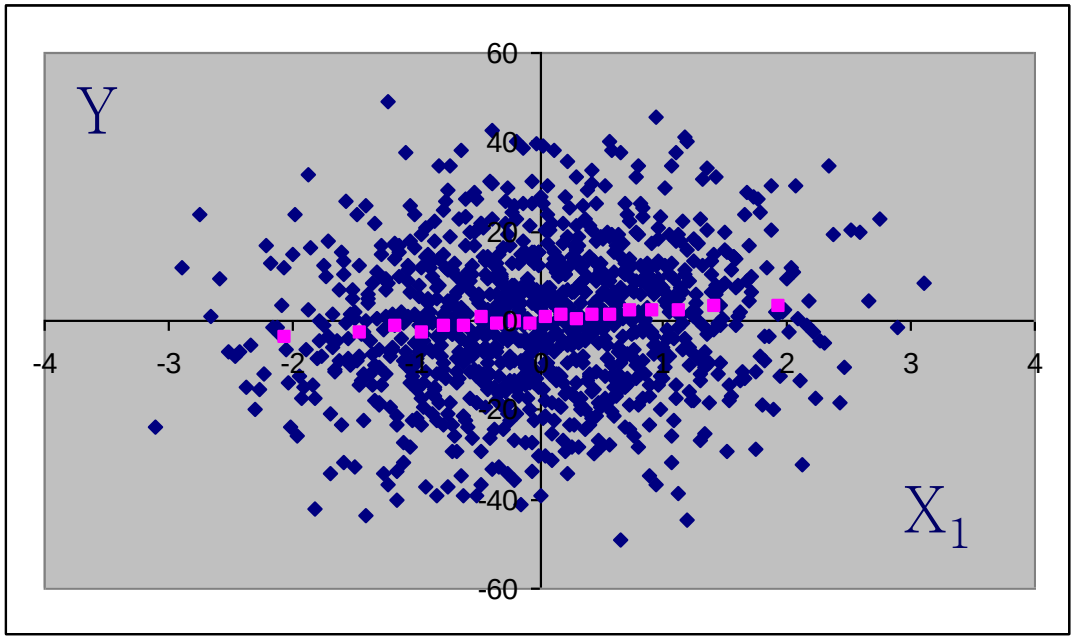
Which factor is more important?

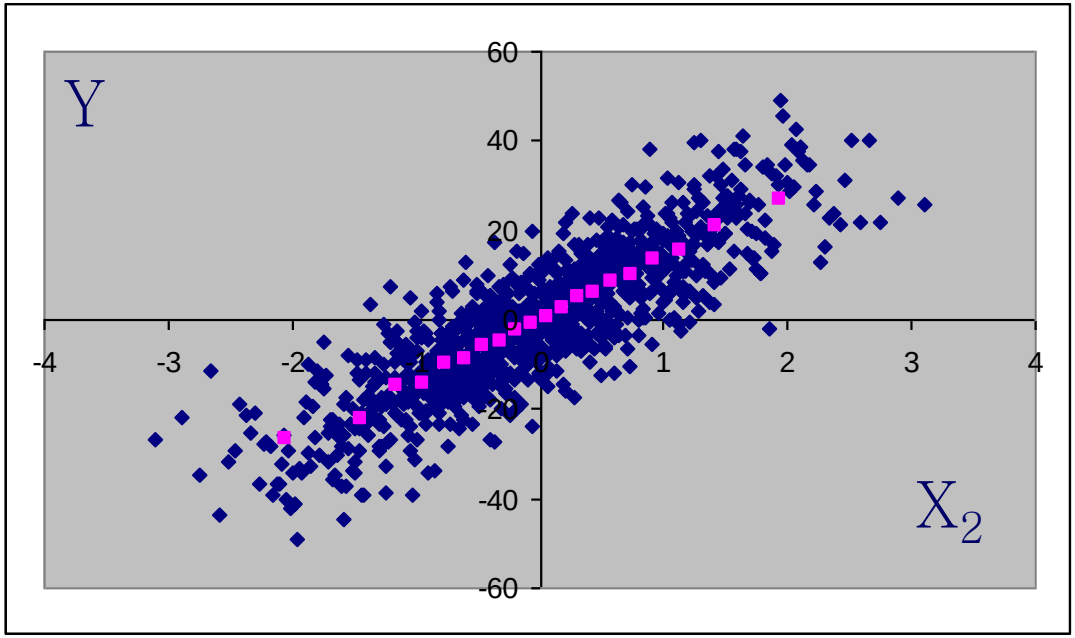




These are ~1,000 points

Divide them in 20 bins of ~ 50 points

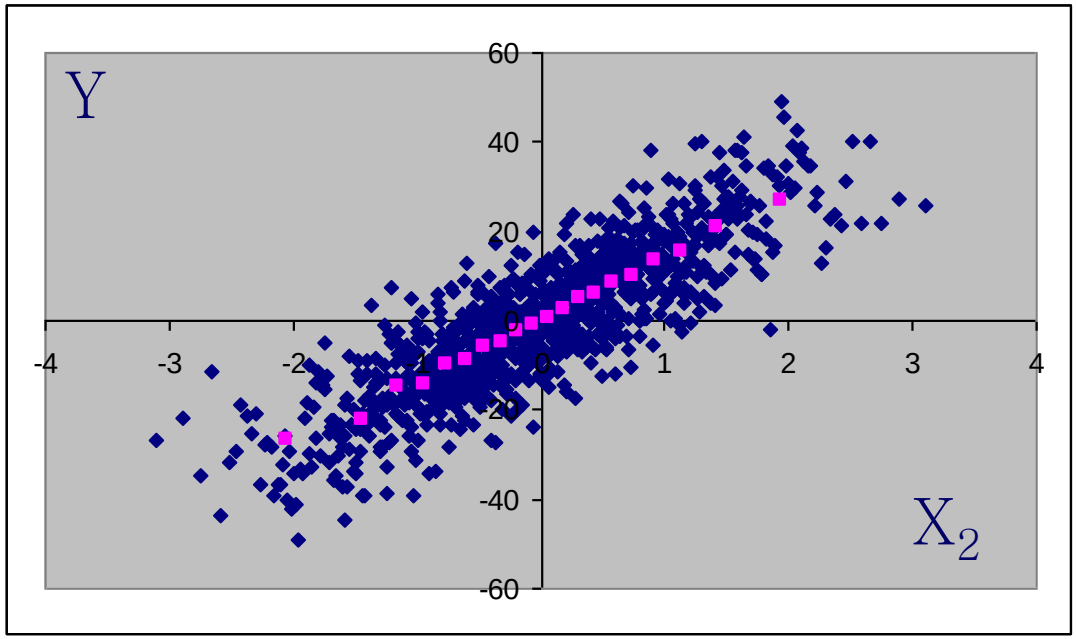




~1,000 blue points

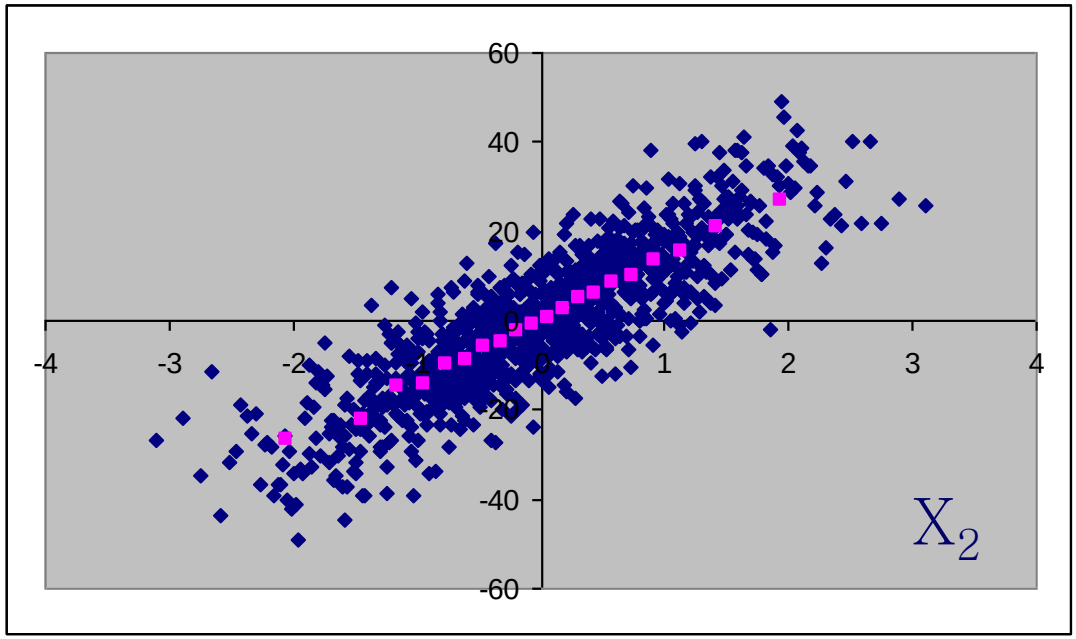
Divide them in 20 bins of ~ 50 points

Compute the bin's average (pink dots)



Each pink point is ~ $E_{\mathbf{x}_{\sim i}}\left(Y \middle| X_i\right)$

Y



Take the variance of the pinkies

$$V_{X_i}\left(E_{\mathbf{X}_{\sim i}}\left(Y|X_i\right)\right)$$

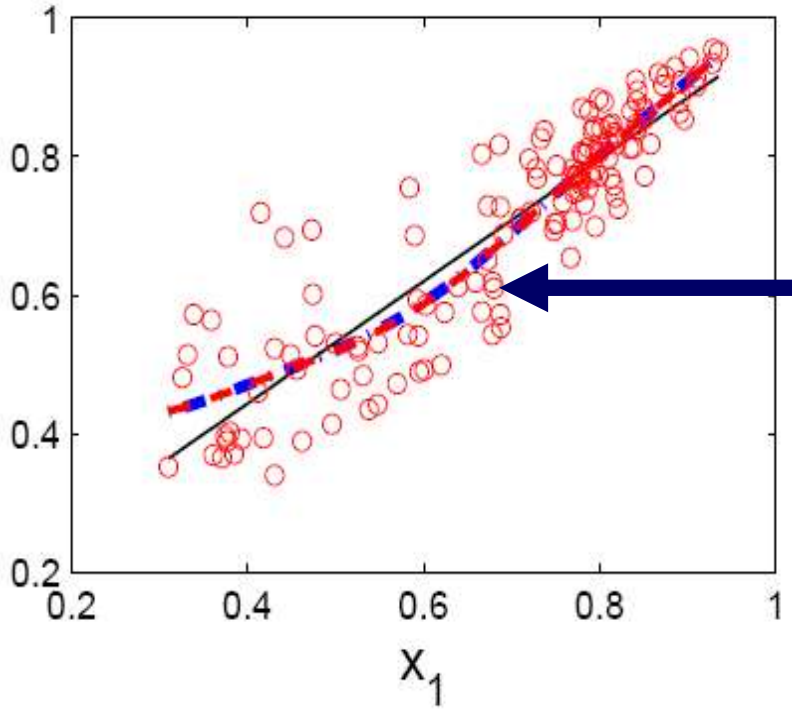
Pearson's correlation ratio

Smoothed curve

$$S_i \equiv \eta_i^2 := \frac{\mathrm{V}_{x_i}\left(\mathrm{E}_{\mathbf{x}_{\sim i}}\left(y \mid x_i\right)\right)}{\mathrm{V}\left(y\right)}$$

First order sensitivity index

Unconditional variance



Smoothed curve

$$\mathrm{E}_{\mathbf{x}_{\sim i}}\left(y \mid x_i\right)$$

First order sensitivity index:

$$\frac{\mathrm{V}_{x_i}\left(\mathrm{E}_{\mathbf{x}_{\sim i}}\left(y \mid x_i\right)\right)}{\mathrm{V}\left(y\right)}$$

$$V_{X_i}\left(E_{\mathbf{X}_{\sim i}}\left(Y|X_i\right)\right)$$

First order effect, or top marginal variance=

= the expected reduction in variance than would be achieved if factor Xi could be fixed.

Why?

Because:

$$V_{X_i}\left(E_{\mathbf{X}_{\sim i}}\left(Y|X_i\right)\right)+$$
$$+E_{X_i}\left(V_{\mathbf{X}_{\sim i}}\left(Y|X_i\right)\right)=V(Y)$$

Easy to prove using $V(Y)=E(Y^2)-E^2(Y)$

Because:

$$V_{X_i}\left(E_{\mathbf{X}_{\sim i}}\left(Y|X_i\right)\right)+$$

$$+E_{X_i}\left(V_{\mathbf{X}_{\sim i}}\left(Y|X_i\right)\right)=V(Y)$$

This is what variance would be left (on average) if Xi could be fixed…

… then this …

$$V_{X_i}\left(E_{\mathbf{X}_{\sim i}}\left(Y|X_i\right)\right)+$$

$$+E_{X_i}\left(V_{\mathbf{X}_{\sim i}}\left(Y|X_i\right)\right)=V(Y)$$

… must be the expected reduction in variance than would be achieved if factor Xi could be fixed

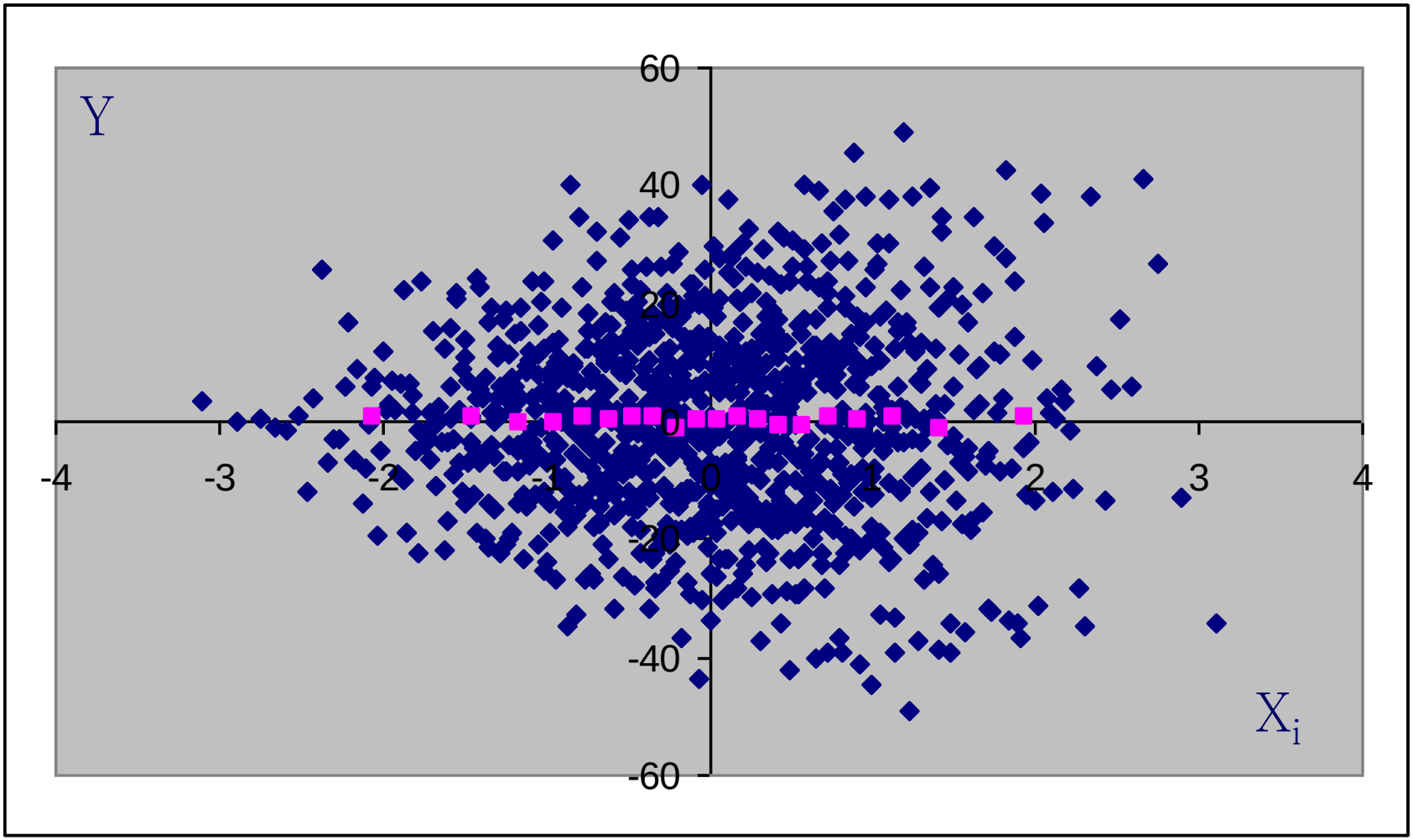
For additive models one can decompose the total variance as a sum of first order effects

$$\sum_i V_{X_i}\left(E_{\mathbf{X}_{\sim i}}\left(Y|X_i\right)\right)\approx V(Y)$$

… which is also how additive models are defined

How about non additive models?

- Is S_i =0?
- Is this factor non-important?



There are terms which capture two-way, three way, … interactions among variables.

All these terms are linked by a formula

Variance decomposition (ANOVA)

$$V(Y) =$$

$$\sum_{i} V_i + \sum_{i,j>i} V_{ij} + \ldots + V_{123..k}$$

Where the last term is an interaction of order k, the number of factors.

Variance decomposition (ANOVA)

$$V_{X_i}\left(E_{\mathbf{X}_{\sim i}}\left(Y\middle|X_i\right)\right)=V_i$$

$$V_{X_iX_j}\left(E_{\mathbf{X}_{\sim ij}}\left(Y\middle|X_iX_j\right)\right)=$$

$$=V_i+V_i+V_{ij}$$

...

Variance decomposition (ANOVA)

When the factors are independent the total variance can be decomposed into main effects and interaction effects up to the order k, the dimensionality of the problem.

Variance decomposition (ANOVA)

When the factors are not independent the decomposition loses its unicity (and hence its appeal)

If fact interactions terms are awkward to handle: second order terms are as many as k(k-1)/2 …

Wouldn't it be handy to have just a single 'importance' terms for all effects, inclusive of first order and interactions?

In fact such terms exist and can be computed easily, without knowledge of the individual interaction terms

Thus given a model $Y=f(X_1,X_2,X_3)$

Instead of

$$V=V_1+V_2+V_3+$$
$$+V_{12}+V_{13}+V_{23}+$$
$$+V_{123}$$

and

$$1=S_1+S_2+S_3+$$
$$+S_{12}+S_{13}+S_{23}+$$
$$+S_{123}$$

We have:

$S_{T1} = S_1 + S_{12} + S_{13} + S_{123}$

(and analogue formulae for S_{T2}, S_{T3}) which can be computed without knowing S_1, S_{12}, S_{13}, S_{123}

S_{T1} is called a total effect sensitivity index

$$E_{\mathbf{X}_{\sim i}}\left(V_{X_i}\left(Y\middle|\mathbf{X}_{\sim i}\right)\right)$$

Total effect, or bottom marginal variance=

= the expected variance than would be left if all factors but Xi could be fixed.

$$\frac{V_{X_i}\left(E_{\mathbf{X}_{\sim i}}\left(Y|X_i\right)\right)}{V(Y)} = S_i$$

$$\frac{E_{\mathbf{X}_{\sim i}}\left(V_{X_i}\left(Y|\mathbf{X}_{\sim i}\right)\right)}{V(Y)} = S_{Ti}$$

Rescaled to [0,1], under the name of first order and total order sensitivity coefficient

Why these measures?

$$V_{X_i}\left(E_{\mathbf{X}_{\sim i}}\left(Y|X_i\right)\right)$$

Factors prioritization

$$E_{\mathbf{X}_{\sim i}}\left(V_{X_i}\left(Y|\mathbf{X}_{\sim i}\right)\right)$$

Fixing (dropping) non important factors

Saltelli A. Tarantola S., 2002, On the relative importance of input factors in mathematical models: safety assessment for nuclear waste disposal, *Journal of American Statistical Association*, 97 (459), 02-709.

Variance based measures are:
-well scaled,
-concise,
-easy to communicate.

Further
- S_i reduces to squared standard regression coefficients for linear model.
- S_{Ti} detect and describe interactions and
- Becomes a screening test at low sample size

(See Campolongo F, Saltelli A, Cariboni, J, 2011, From screening to quantitative sensitivity analysis. A unified approach, *Computer Physics Communication*, 182 (4), pp. 978-988.)

Both indices can be computed via Monte Carlo

We use quasi random sequences developed by I.M. Sobol'

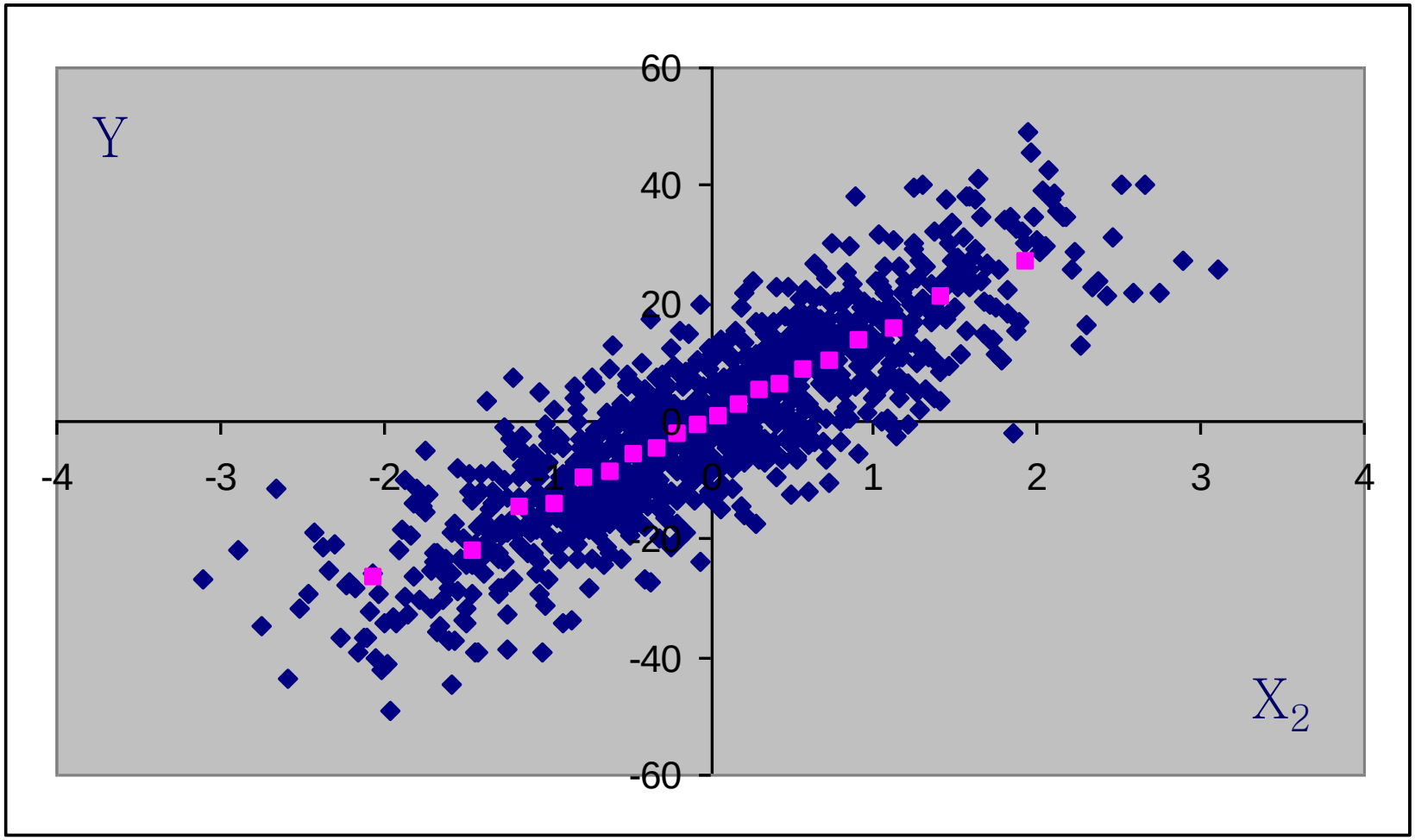
Estimation procedures:

- No brute force. A double loop is not needed, though the measures are expresses as V(E(•)) and E(V(•)).
- For S_i quick estimation procedures are available which are k-independent.
- For S_{Ti} estimation procedures are mostly k-dependent (unless … active area of research…).

Summary for variance based measures:

1. Easy-to-code, Monte Carlo – better on quasi-random points. Estimate of the error available.

2. The main effect can be made cheap; its computational cost does not depend upon k.



Easy to smooth and interpolate!

Summary for variance based measures:

3. The total effect is more expensive; its computational cost is $(k+1)N$ where N is one of the order of one thousand (unless e.g. using emulators …).

Things to keep in mind for a good sensitivity analysis

Why SA? What is the question?

Sensitivity analysis is not “run” on a model but on a model once applied to a case.

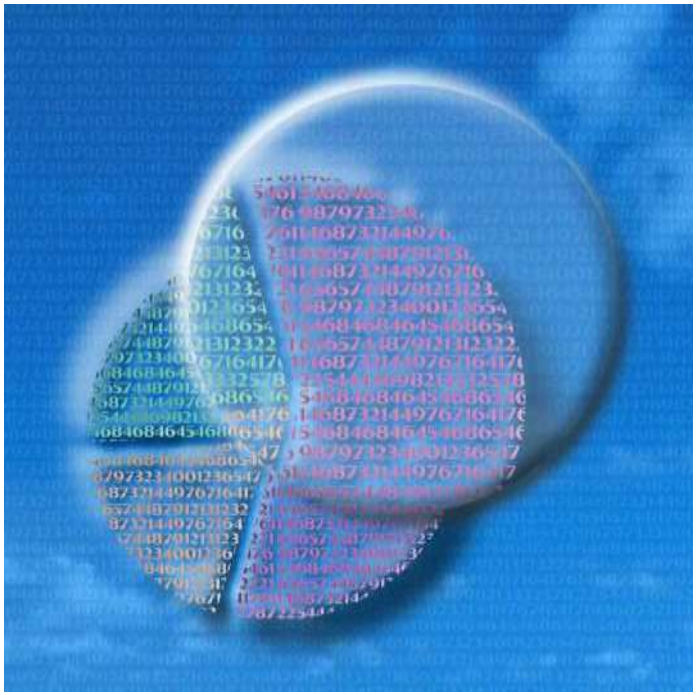
Sensitivity analysis should not be used to hide assumptions.

SA for confirmation or for falsification? The latter works better.

If SA shows that a question cannot be answered by the model then find another question/model which can be treated meaningfully.

Discussion point

- Why doing a sensitivity analysis if it can undermine an laborious quantification exercise?
- What do I do if this happens to be the case?



END

UNIVERSITÉ DE LA REUNION
S'ouvrir aux mondes

SAMO 2016

Tapez vos mots-clé ici...

HOME | ABSTRACT SUBMISSION | REGISTRATION | ACCOMODATION | VISITOR INFORMATION | YOUR TRAVEL

Eighth International Conference on Sensitivity Analysis of Model Output - November 30th to December 3rd 2016, Reunion Island

A Conference celebrating the 90th birthday of Ilya M. Sobol'

SAMO 2016 (Reunion, France)